

## Fig. 6A

C.anc.env (subtype C ancestral env. The amino acid sequence is different from Los Alamos Database August 2002)

GCCGCCATGCGCGTGATGGGCATCCTGCGCAACTGCCAGCAGTGGTGAT  
CTGGGGCATCCTGGGCTTCTGGATGCTGATGATCTGCTCCGTGGTGGGCA  
ACCTGTGGGTGACCGTGTACTACGGCGTGC CCGTGTGGAAAGGAGGCCAAG  
ACCA CCTGTCTGCGCCTCCGACG CCAAGGCC TA CGAGCGCGAGGTGCA  
CAACGTGTGGGCCACCCACGCCTGCGTGCC CACCGACCCCAA CCC CCAGG  
AGATGGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACGAC  
ATGGTGGAC CAGATGCACGAGGACATCA TCTCCCTGTGGGAC CAGTCCCT  
GAAGCCCTGCGTGAAGCTGACCCCTGTGCGTGACCTGAACTGCA CCA  
ACGTGACCAACGCCACCAACAACCTACAACCGGC GAGATGAAGAACTGC  
TCTTCAACATCACCA CCGAGCTGCGCGACAAGAA GAAGAAGGAGTACGC  
CCTGTTCTACCGCCTGGACATCGTGCCCCTGAACGAGAACTCCTCGAGT  
ACCGCCTGATCAACTGCAACACCTCCGCATCA CCGAGGCCTGCCCAAAG  
GTGTCCTTCGACCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTACGC  
CATCCTGAAGTGCAACAA CAAGACCTTCAA CGGCA CCGGCCCTGCAACA  
ACGTGTCCA CCGTG CAGTGCA CCGCATCAAG CCGGTGGTGTCCACC  
CAGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGGAGATCATCATCCTC  
CGAGAACCTGACCGACAACGCAAGACCATCATCGTG CAGCTGAACGAGT  
CCGTGGAGATCTGTGTG CACCCGCCCAA CAACAACACCCGCAAGTCCATG  
CGCATCGGCCCCCGGCCAGACCTTCTACGCCACCGGCGACATCATCGGCGA  
CATCCGCCAGGCCACTGCAACATCTCCGAGGACAAGTGGAA CAAGACCC  
TGCAGCAGGTGGCCGAGAAGCTGGGCAAGCACTTCCCAA CAAGACCATC  
ACCTTCGAGCCCTCCTCCGGCGGCGACCTGGAGATCACCA CCACTCCTT  
CAACTGCCGCGCGAGTTCTTCTACTGCAA CACCTCCAAGCTGTTCAACT  
CCACCTACAACAACAACA CCAACTCCA CTCCACCATCACCTGCCCCTGC  
CGCATCAAGCAGATCATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTA  
CGCCCCCCCCCATCGCCGGCAA CATACCTGCAAGTCCAACATCACCGGCC  
TGCTGCTGACCCGCGACGCGCGCAAGGAGAACA CCAACCGAGACCTTCGC  
CCCGGCCGGCGGCGACATGCGCGACA ACTGGCGCTCCGAGCTGTACAA GTA  
CAAGGTGGTGGAGATCAAGCCCTGGGCGTGGCCCCACCGAGGGCAAGC  
GCCGCGTGGTGGAGCGCGAGAAGCGCGCGTGGGCTGGGCGCCGTGTTCT  
CTGGGCTTCTGGGCGCCGCCGGCTCCA CCAATGGGCGCCGCTCCATCAC  
CCTGACCGTG CAGGCCCGCCAGCTGCTGTC CGGCATCGTG CAGCAGCAGT  
CCAACCTGCTGCGCGCATCGAGGCCAGCAGCACATGCTGCAGCTGACC  
GTGTGGGGCATCAAGCAGCTG CAGGCCCGCTGCTGGCCATGGAGCGCTA  
CCTGAAGGACCAGCAGCTGCTGGGCATCTG GGGCTGCTCCGGCAAGCTGA  
TCTGCACCA CCGCGTGCCCTGGAACTCCTCCTGGTCCAA CAAGTCCCTG  
GACGACATCTGGGACAACATGACCTGGA TGGAGTGGGACCGCGAGATCTC  
CAACTACACCGACACCATCTACCGCCTGCTGGAGGAGTCC CAGAAC CAGC  
AGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGACTCCTGGGAGAACCTG  
TGGAACCTGGTT CGACATCA CCAACTGGCTGTGGTACATCAAGATCTT CAT  
CATGATCGTGGGCGGCCTGATCGGCCTGCGCATCATCTTCCGCTGCTGT  
CCATCGTGAA CCGCGTGCGCCAGGGCTA CTCCCCCTGTCTCTCCAGACC  
CTGACCCCCAA CCCCCGCGGCCCGACCGCCTGGAGCGCATCGAGGAGGA  
GGGCGGCGAGCAGGACCGCGACCGCTCCATCCGCTGGTGTCGGCTTCC  
TGGCCCTGGCCTGGGACGACCTGCGCTCCCTGTGCTGTTCTCCTAC CAC  
CGCTGCGCGA CTTCA TCTGATCGCCGCCGCGACCGTGGAGCTGCTGGG  
CCGCTCCTCCCTGCGCGCCTGCAGCGCGCTGGGAGGCCCTGAAGTACC  
TGGGCTCCCTGGTG CAGTACTGGGGCCAGGAGCTGAAGAA GTCCGCCATC  
TCCCTGCTGGA CACCATCGCCATCGCCGTGGCCGAGGGCA CCGACCGCAT  
CATCGAGGTGGTGAGCGCGCCTGCCGCGCATCTGAAACATCCCCCGCC  
GCATCCGCCAGGGCTTCGAGGCCGCCCTGCTGTAA

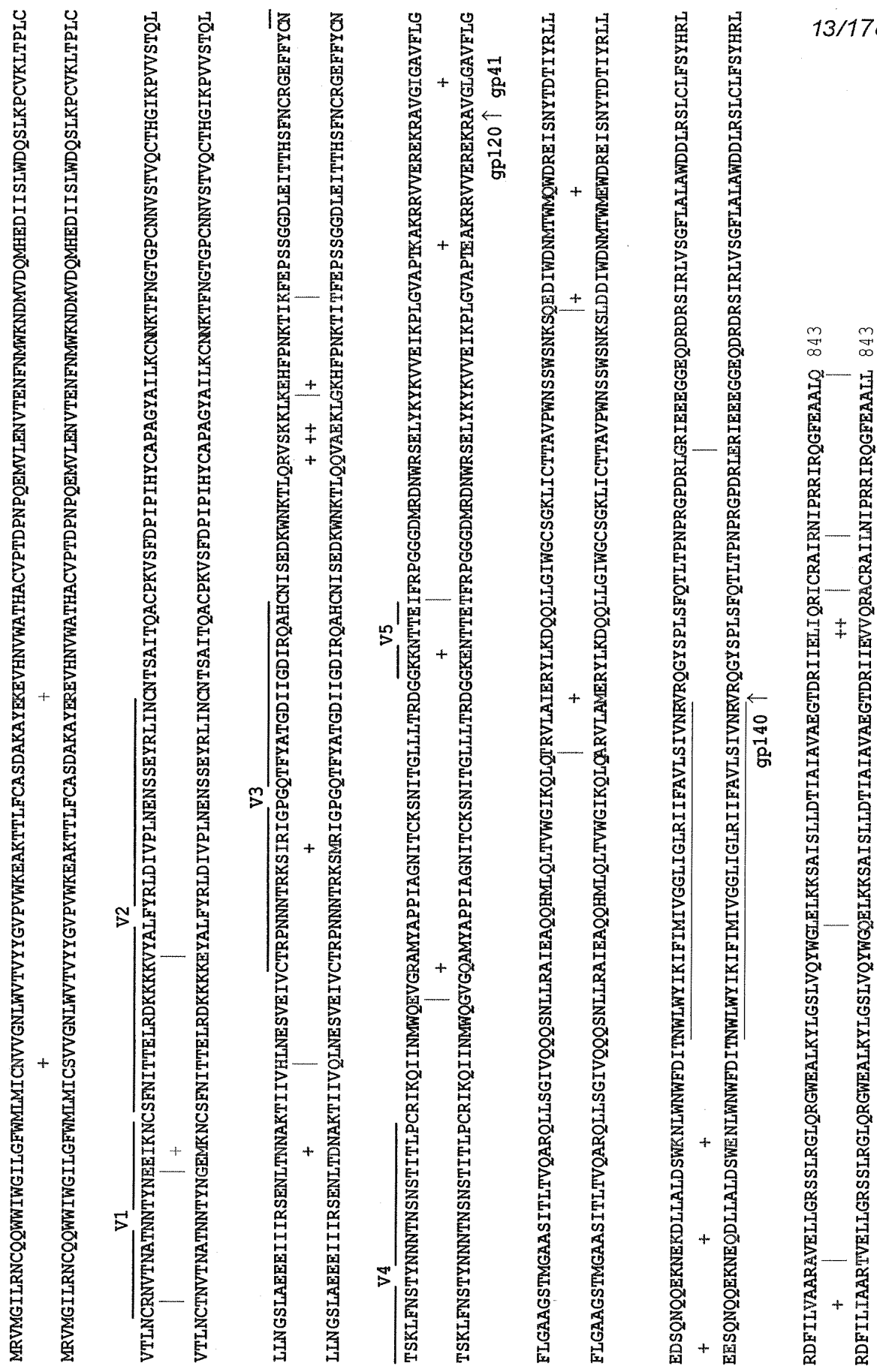
9/178

## Fig. 6B

C.con.env (subtype C consensus env. The amino acid sequence is different from Los Alamos Database August 2002)

GCCGCCATGCGCGTGATGGGCATCCTGCGCAACTGCCAGCAGTGGTGGAT  
CTGGGGCATCCTGGGCTTCTGGATGCTGATGATCTGCAACGTGGTGGGCA  
ACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAAG  
ACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGAGAAGG AGGTGCA  
CAACGTGTGGGGCCACCCACGCCTGCGTGCCCAACCGACCCCAACCCCCAGG  
AGATGGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACGAC  
ATGGTGGACCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTCCCT  
GAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGAACTGCCGCA  
ACGTGACCAACGCCACCAACAACACCTACAACGAGGAGATCAAG AACTGC  
TCCTTCAACATCACCAACCGAGCTGCGCGACAAGAAGAAGGTGTACGC  
CCTGTTCTACCGCCTGGACATCGTGCCCCCTGAACGAGAACTCCTCCGAGT  
ACCGCCTGATCAACTGCAACACCTCCGCCATCACCCAGGCCTGCCCAAG  
GTGTCTTCGACCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTACGC  
CATCCTGAAGTGCAACAACAAGACCTTCAACGGCACCCGGCCCCCTG CAACA  
ACGTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGTCCACC  
CAGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGGAGATCATCATCCGCTC  
CGAGAACCTGACCAACAACGCCAAGACCATCATCGTGACCTGAACGAGT  
CCGTGGAGATCGTGTGCACCCGCCCAACAACAACACCCGCAAGTCCATC  
CGCATCGGCCCCGGCCAGACCTTCTACGCCACCGGCGACATCATCG GCGA  
CATCCGCCAGGCCCCACTGCAACATCTCCGAGGACAAGTGGAAACAAGACCC  
TGCAGCGCGTGTCCAAGAAGCTGAAGGAGCACTTCCCCAACAAGACCATC  
AAGTTCGAGCCCTCCTCCGGCGGCGACCTGGAGATCACCAACCCACTCCTT  
CAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCAAGCTGTTCAACT  
CCACCTACAACAACAACCAACTCCAACCTCCACCATCACCTTGCCC TGC  
CGCATCAAGCAGATCATCAACATGTGGCAGGAGGTGGGCCGCGCCATGTA  
CGCCCCCCCCATCGCCGGCAACATACCTGCAAGTCCAACATCACCGGCC  
TGCTGCTGACCCGCGACGGCGGCAAGAAGAACACCACCGAGATCTTCCGC  
CCCCGGCGGCGGCGACATGCGCGACAACCTGGCGCTCCGAGCTGTACAAGTA  
CAAGGTGGTGGAGATCAAGCCCCCTGGGCGTGGCCCCCACCAGGCCAA GC  
GCCGCGTGGTGGAGCGCGAGAAGCGCGCCGTGGGCATCGGCGCCGTGTTT  
CTGGGCTTCTGGGCGCCGCGGCTCCACCATGGGCGCCGCTCCATCAC  
CCTGACCGTGCAGGCCCCGCCAGCTGCTGTCCGGCATCGTGCAGCAGCAGT  
CCAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACATGCTGCAGCTGACC  
GTGTGGGGCATCAAGCAGCTGCAGACCCGCGTGTGGCCATCGAGCGCTA  
CCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGAAGCTGA  
TCTGCACCAACCGCCGTGCCCTGGAACCTCCTCCTGGTCCAACAAGTCCCAG  
GAGGACATCTGGGACAACATGACCTGGATGCAGTGGGACCGCGAGATCTC  
CAACTACACCGACACCATCTACCGCCTGCTGGAGGACTCCCAGAACCAGC  
AGGAGAAGAACGAGAAGGACCTGCTGGCCCTGGACTCCTGGAAGAACCTG  
TGGAACCTGGTTCGACATCACCAACTGGCTGTGGTACATCAAGATCTTCAT  
CATGATCGTGGGCGGCCTGATCGGCCTGCGCATCATCTTCGCCGTGCTGT  
CCATCGTGAACCGCGTGCGCCAGGGCTACTCCCCCTGTCTTCCAGACC  
CTGACCCCCAACCCCCGCGCCCCGACCGCCTGGGCCGCATCGAGGAGGA  
GGGCGGCGAGCAGGACCGCGACCGCTCCATCCGCCTGGTGTCCGGCTTCC  
TGGCCCTGGCCTGGGACGACCTGCGCTCCCTGTGCCTGTTCTCCTACCAC  
CGCCTGCGCGACTTCATCCTGGTGGCCGCCCGCGCCGTGGAGCTGCTGGG  
CCGCTCCTCCCTGCGCGGCTGCAGCGCGGCTGGGAGGCCCTGAAGTACC  
TGGGCTCCCTGGTGCAGTACTGGGGCCTGGAGCTGAAGAAGTCCGCCATC  
TCCCTGCTGGACACCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCAT  
CATCGAGCTGATCCAGCGCATCTGCCGCGCCATCCGCAACATCCCCCGCC  
GCATCCGCCAGGGCTTCGAGGCGGCCCTGCAGTAA

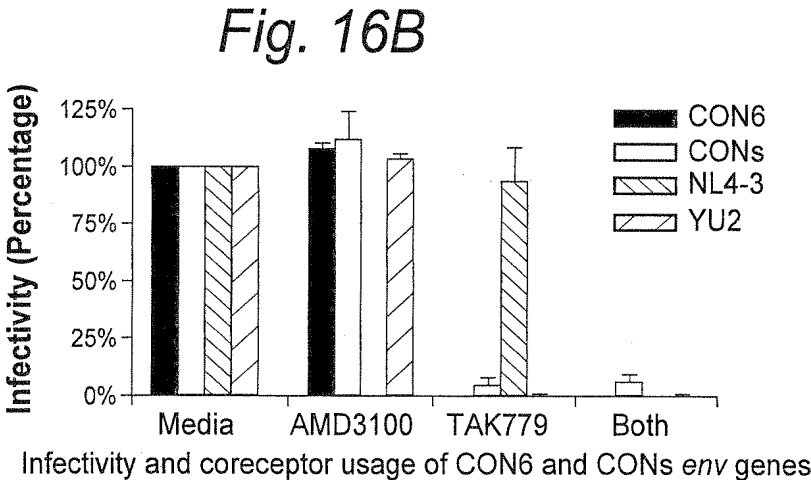
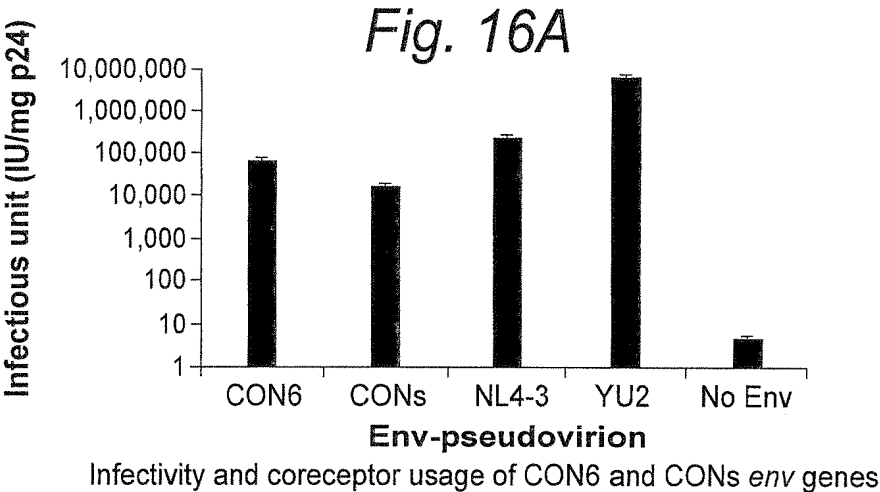
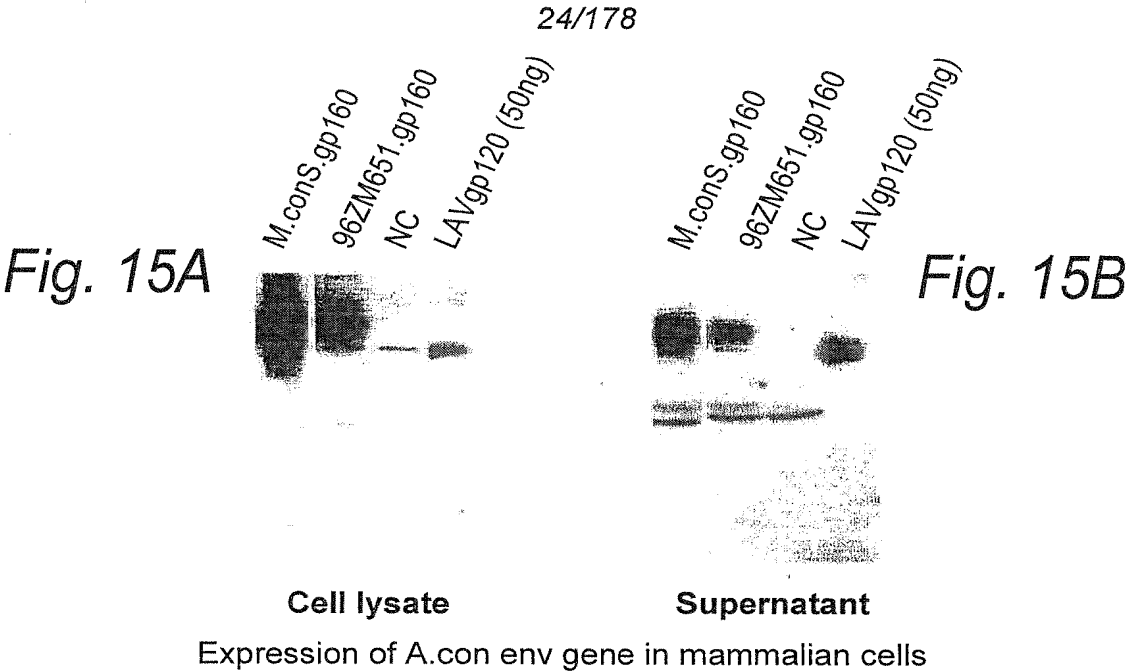
Fig. 8



## Fig. 14B

CONs.env (gorup M consensus env gene. This one contain the consensus sequence for variable regions in env gene. The identical amino acid sequences as in the public domain)

```
GCCGCCGCCATGCGCGTGCGCGGCATCCAGCGCAACTGCCAGCACCTGTG
GCGCTGGGGCACCCTGATCCTGGGCATGCTGATGATCTGCTCCGCCGCCG
AGAACCTGTGGGTGACCGTGTAACGCGTACGGCGTGCCCGTGTGGAAGGAGGCC
AACACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCTGCGTGCCACCGACCCCAACCCCC
AGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAAC
AACATGGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTC
CCTGAAGCCCTGCGTGAAAGCTGACCCCCCTGTGCGTGACCCTGAACTGCA
CCAACGTGAACGTGACCAACACCACCAACAACACCGAGGAGAAGGGCGAG
ATCAAGAACTGCTCCTTCAACATCACCACCGAGATCCGCGACAAGAAGCA
GAAGGTGTACGCCCTGTTCTACCGCCTGGACGTGGTGGCCATCGACGACA
ACAACAACAACCTCCTCCAACCTACCGCCTGATCAACTGCAACACCTCCGCC
ATCACCCAGGCCTGCCCAAGGTGTCTTTCGAGCCCATCCCCATCCACTA
CTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCA
ACGGCACCGGCCCTGCAAGAACGTGTCCACCGTGACGTGCACCCACGGC
ATCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGCTCCCTGGCCGA
GGAGGAGATCATCATCCGCTCCGAGAACATCACCACAACGCCAAGACCA
TCATCGTGACGTGAACGAGTCCGTGGAGATCAACTGCACCCGCCCAAC
AACAACACCCGCAAGTCCATCCGCATCGGCCCGGCCAGGCCTTCTACGC
CACCGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCTCCG
GCACCAAGTGGAACAAGACCCTGCAGCAGGTGGCCAAGAAGCTGCGCGAG
CACTTCAACAACAAGACCATCATCTTCAAGCCCTCCTCCGGCGGCGACCT
GGAGATCACCACCCACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGCA
ACACCTCCGGCCTGTTCAACTCCACCTGGATCGGCAACGGCACCAAGAAC
AACAACAACACCAACGACACCATCACCCTGCCCTGCCGCATCAAGCAGAT
CATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTACGCCCCCCCCCATCG
AGGGCAAGATCACCTGCAAGTCCAACATCACCAGCCTGCTGCTGACCCGC
GACGGCGGCAACAACAACACCAACGAGACCGAGATCTTCCGCCCGGCGG
CGGCGACATGCGCGACAACCTGGCGCTCCGAGCTGTACAAGTACAAGGTGG
TGAAGATCGAGCCCCTGGGCGTGGCCCCCAACAAGGCCAAGCGCCGCGTG
GTGGAGCGCGAGAAGCGCGCCGTGGGCATCGGCGCCCGTGTTCCTGGGCTT
CCTGGGCGCCGCGGCTCCACCATGGGCGCCGCTCCATCACCTGACCG
TGCAGGCCCCGCCAGCTGCTGTCCGGCATCGTGCAGCAGCAGTCCAACCTG
CTGCGCGCCATCGAGGCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGG
CATCAAGCAGCTGCAGGCCCGCGTGTGCTGGCCGTGGAGCGCTACCTGAAGG
ACCAGCAGCTGCTGGGCATCTGGGCGTGTCCGGCAAGCTGATCTGCACC
ACCACCGTGCCCTGGAACCTCCTCCTGGTCCAACAAGTCCCAGGACGAGAT
CTGGGACAACATGACCTGGATGGAGTGGGAGCGCGAGATCAACAACATA
CCGACATCATCTACTCCCTGATCGAGGAGTCCCAGAACCAGCAGGAGAAG
AACGAGCAGGAGCTGCTGGCCCTGGACAAGTGGGCCTCCCTGTGGAACCTG
GTTTCGACATCACCAACTGGCTGTGGTACATCAAGATCTTCATCATGATCG
TGGGCGGCCTGATCGGCCTGCGCATCGTGTTTCGCCGTGCTGTCCATCGTG
AACC GCGTGCGCCAGGGCTACTCCCCCTGTCTTCCAGACCCTGATCCC
CAACCCCCGCGGCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGCGGCG
AGCAGGACCGCGACCGCTCCATCCGCCTGGTGAACGGCTTCCTGGCCCTG
GCCTGGGACGACCTGCGCTCCCTGTGCTGTTCCTACCACCGCCTGCG
CGACTTCATCCTGATCGCCGCCCGCACCGTGGAGCTGCTGGGCCGCAAGG
GCCTGCGCCGCGGCTGGGAGGCCCTGAAGTACCTGTGGAACCTGCTGCAG
TACTGGGGCCAGGAGCTGAAGAACTCCGCCATCTCCCTGCTGGACACCAC
CGCCATCGCCGTGGCCGAGGGCACCGACCGCGTGATCGAGGTGGTGCAGC
GCGCCTGCCGCGCCATCCTGAACATCCCCCGCCGCATCCGCCAGGGCCTG
GAGCGCGCCCTGCTGTAA
```

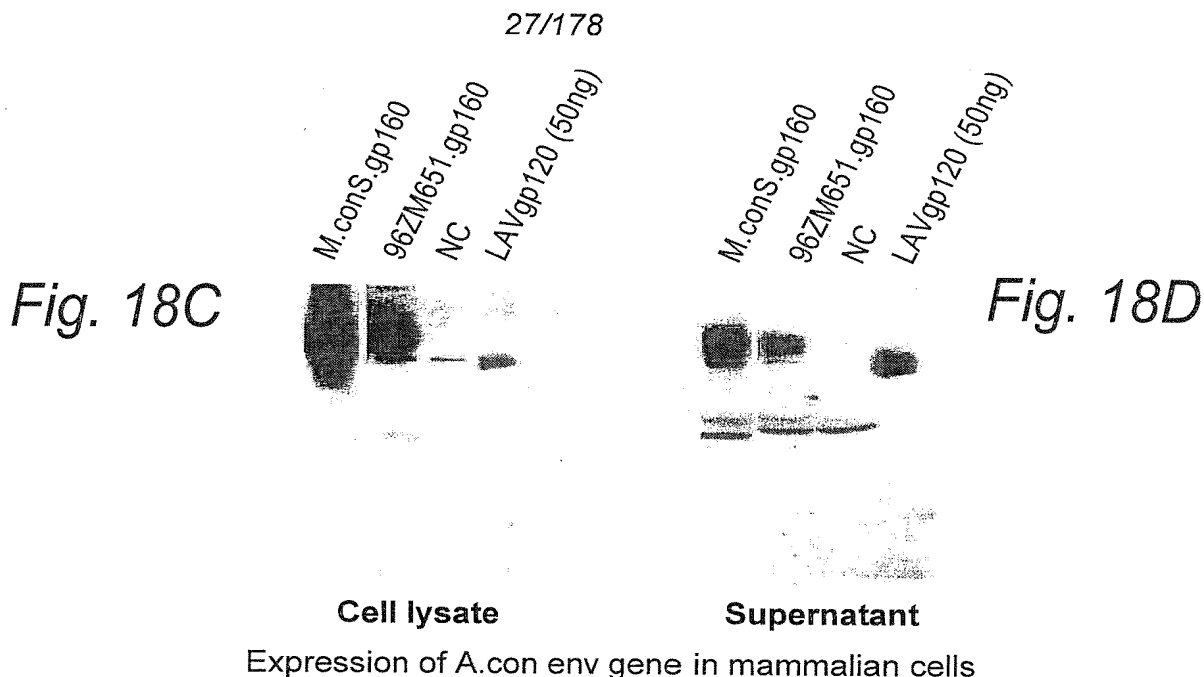


26/178

## Fig. 18B

A.con.env (subtype A consensus env. Identical amino acid sequence to that in the public domain)

GCCGCCGCCATGCGCGTGATGGGCATCCAGCGCAACTGCCAGCACCTGTG  
GCGCTGGGGCACCATGATCCTGGGCATGATCATCATCTGCTCCGCCGCCG  
AGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGACGCC  
GAGACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGACACCGAGGT  
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCC  
AGGAGATCAACCTGGAGAACGTGACCGAGGAGTTCAACATGTGGAAGAAC  
AACATGGTGGAGCAGATGCACACCGACATCATCTCCCTGTGGGACCAGTC  
CCTGAAGCCCTGCGTGGAAGCTGACCCCCCTGTGCGTGACCCCTGAACTGCT  
CCAACGTGAACGTGACCACCAACATCACCAACATCACCGACAACATGAAG  
GGCGAGATCAAGAACTGCTCCTTCAACATGACCACCGAGCTGCGCGACAA  
GAAGCAGAAGGTGTACTCCCTGTTCTACAAGCTGGACGTGGTGCAGATCA  
ACAAGTCCAACCTCCTCCTCCAGTACCGCCTGATCAACTGCAACACCTCC  
GCCATCACCCAGGCCGTGCCCAAGGTGTCTTTCGAGCCCATCCCCATCCA  
CTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAAGGACAAGGAGT  
TCAACGGCACC GGCCCCCTGCAAGAACGTGTCCACCGTGCAAGTGACCCAC  
GGCATCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGCTCCCTGGC  
CGAGGAGGAGGTGATGATCCGCTCCGAGAACATCACCAACAACGCCAAGA  
ACATCATCGTGCAGCTGACCAAGCCCGTGAAGATCAACTGCACCCGCCCC  
AACAACAACACCCGCAAGTCCATCCGCATCGGCCCGGCCAGGCCTTCTA  
CGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACGTGT  
CCCGCACCGAGTGGAAACGAGACCCCTGCAGAAGGTGGCCAAGCAGCTGCGC  
AAGTACTTCAACAACAAGACCATCATCTTCAACCAACTCCTCCGCGCGCGA  
CCTGGAGATCACCAACCCACTCCTTCAACTGCGGCGGCGAGTTCTTCTACT  
GCAACACCTCCGGCCTGTTCAACTCCACCTGGAACGGCAACGGCACCAAG  
AAGAAGAACTCCACCGAGTCCAACGACACCATCACCCCTGCCCTGCCGCAT  
CAAGCAGATCATCAACATGTGGCAGCGCGTGGGCCAGGCCATGTACGCCC  
CCCCCATCCAGGGCGTGATCCGCTGCGAGTCCAACATCACCGGCCTGCTG  
CTGACCCGCGACGGCGGC GACAACAACCTCCAAGAACGAGACCTTCCGCCC  
CGGCGGCGGCGACATGCGCGACAACCTGGCGCTCCGAGCTGTACAAGTACA  
AGGTGGTGAAGATCGAGCCCCCTGGGCGTGGCCCCCACCAGGCCAAGCGC  
CGCGTGGTGGAGCGCGAGAAGCGCGCCGTGGGCATCGGCGCCGTGTTCTCT  
GGGCTTCCCTGGGCGCCGCGGCTCCACCATGGGCGCCGCTCCATCACCC  
TGACCGTGACAGGCCCGCCA GCTGCTGTCCGGCATCGTGACAGCAGTCC  
AACCTGCTGCGCGCCATCGAGGCCAGCAGCACCTGCTGAAGCTGACCGT  
GTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACC  
TGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGCAAGCTGATC  
TGCACCACCAACGTGCCCTGGAACCTCCTCCTGGTCCAACAAGTCCCAGTC  
CGAGATCTGGGACAACATGA CCTGGCTGCAGTGGGACAAGGAGATCTCCA  
ACTACACCGACATCATCTACAACCTGATCGAGGAGTCCCAGAACCGACAG  
GAGAAGAACGAGCAGGACCTGCTGGCCCTGGACAAGTGGGCCAACCTGTG  
GAACTGGTTCGACATCTCCAACCTGGCTGTGGTACATCAAGATCTTCATCA  
TGATCGTGGGCGGCCTGATCGGCCTGCGCATCGTGTTCCGCCGTGCTGTCC  
GTGATCAACCGCGTGCGCCAG GGCTACTCCCCCTGTCTTCCAGACCCA  
CACCCCCAACCCCGGCGGCCTGGACCGCCCCGGCCGCATCGAGGAGGAGG  
GCGGCGAGCAGGGCCGCGACCGCTCCATCCGCCTGGTGTCCGGCTTCTTG  
GCCCTGGCCTGGGACGACCTGCGCTCCCTGTGCCTGTTCTCCTACCACCG  
CCTGCGCGACTTCATCCTGATCGCCGCCCGCACCGTGGAGCTGCTGGGCC  
ACTCCTCCCTGAAGGGCCTGCG CCTGGGCTGGGAGGGCCTGAAGTACCTG  
TGGAACCTGCTGCTGTACTGGGGCCGCGAGCTGAAGATCTCCGCCATCAA  
CCTGCTGGACACCATCGCCATCGCCGTGGCCGGCTGGACCGACCGCGTGA  
TCGAGATCGGCCAGCGCATCTGCCGCGCCATCCTGAACATCCCCCGCCG  
ATCCGCCAGGGCCTGGAGCGCGCCCTGCTGTAA



*Fig. 19A*

M.con.gag (group M consensus gag. Identical amino acid sequence to that in the public domain)

```

GCCGCCGCCATGGGCGCCCGCGCCTCCGTGCTGTCCGGCGGCAAGCTGGA
CGCCTGGGAGAAGATCCGCCTGCGCCCCGGCGGCAAGAAGAAGTACCGCC
TGAAGCACCTGGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCCTGAAC
CCCGGCCTGCTGGAGACCTCCGAGGGCTGCAAGCAGATCATCGGCCAGCT
GCAGCCCGCCCTGCAGACCGGCTCCGAGGAGCTGCGCTCCCTGTACAACA
CCGTGGCCACCCTGTACTGCGTGACACGCGCATCGAGGTGAAGGACACC
AAGGAGGCCCTGGAGAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAA
GACCCAGCAGGCCGCGCCGACAAAGGGCAACTCCTCCAAGGTGTCCCAGA
ACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACACGAGCCATC
TCCCCCGCACCTGAACGCTGGGTGAAGGTGATCGAGGAGAAGGCCTT
CTCCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCC
CCCAGGACCTGAACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCC
ATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCG
CCTGCACCCCGTGACGCCGGCCCCATCCCCCGGCCAGATGCGCGAGC
CCCGCGGCTCCGACATCGCCGGCACACCTCCACCCTGCAGGAGCAGATC
GCCTGGATGACCTCCAACCCCCCATCCCCGTGGGCGAGATCTACAAGCG
CTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCCCCCGTGT
CCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTG
GACCGCTTCTTCAAGACCTGCGCGCCGAGCAGGCCACCCAGGACGTGAA
GAACTGGATGACCGACACCCTGCTGGTGCAGAACGCCAACCCGACTGCA
AGACCATCCTGAAGGCCCTGGGCCCCGGCGCCACCTGGAGGAGATGATG
ACCGCCTGCCAGGGCGTGGGCGGCCCGGCCACAAGGCCCGCGTGTGGC
CGAGGCCATGTCCCAGGTGACCAACGCCGCCATCATGATGCAGCGCGGCA
ACTTCAAGGGCCAGCGCCGATCATCAAGTGCTTCAAGTGGCGCAAGGAG
GGCCACATCGCCCCGCAACTGCCGCGCCCCCGCAAGAAGGGCTGTGGAA
GTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCA
ACTTCTGGGCAAGATCTGGCCCTCCAACAAGGGCCGCCCGGCCAACTTC
CTGCAGTCCCGCCCCGAGCCACCGCCCCCCCCCGCCGAGTCCTTCGGCTT
CGGCGAGGAGATCACCCCTCCCCCAAGCAGGAGCCCAAGGACAAGGAGC
CCCCCTGACCTCCCTGAAGTCCCTGTTTCGGCAACGACCCCTGTCCAG
TAA
    
```

M.con.pol.nuc

28/178

Fig. 19B

GCCGCCGCCATGCCCCAGATCACCTGTGGCAGCGCCCCCTGGTGACCAT  
CAAGATCGGCGGCCAGCTGAAGGAGGCCCTGCTGGCCACCGGCGCCGACG  
ACACCGTGCTGGAGGAGATCAACCTGCCCGGCAAGTGGAAGCCCAAGATG  
ATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCT  
GATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGGTGGGCCCCA  
CCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGATCGGCTGCACC  
CTGAACTTCCCCATCTCCCCATCGAGACCGTGCCCCGTGAAGCTGAAGCC  
CGGCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCCTGACCGAGGAGAAGA  
TCAAGGCCCTGACCGAGATCTGCACCGAGATGGAGAAGGAGGGCAAGATC  
TCCAAGATCGGCCCCGAGAACCCCTACAACACCCCCATCTTCGCCATCAA  
GAAGAAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA  
ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCC  
GCCGGCTGAAGAAGAAGTCCGTGACCGTGCTGGACGTGGGCGGACGC  
CTACTTCTCCGTGCCCCCTGGACGAGGACTTCCGCAAGTACACCGCCTTCA  
CCATCCCCCTCCATCAACAACGAGACCCCCGGCATCCGCTACCAGTACAAC  
GTGCTGCCCCAGGGCTGGAAGGGCTCCCCCGCCATCTTCAGTCCTCCAT  
GACCAAGATCCTGGAGCCCTTCCGCACCCAGAACCCCGAGATCGTGATCT  
ACCAGTACATGGACGACCTGTACGTGGGCTCCGACCTGGAGATCGGCCAG  
CACCGCGCCAAGATCGAGGAGCTGCGCGAGCACCTGCTGCGCTGGGGCTT  
CACCACCCCCGACAAGAAGCACCAAGGAGCCCCCTTCCTGTGGATGG  
GCTACGAGCTGCACCCCGACAAGTGGACCGTGACGCCATCCAGCTGCCC  
GAGAAGGACTCCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCT  
GAACTGGGCCTCCAGATCTACCCCGGCATCAAGGTGAAGCAGCTGTGCA  
AGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAG  
GAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGAAGGAGCCCGT  
GCACGGCGTGTACTACGACCCCTCCAAGGACCTGATCGCCGAGATCCAGA  
AGCAGGGCCAGGACCAGTGGACCTACCAGATCTACCAGGAGCCCTTCAAG  
AACCTCAAGACCGGCAAGTACGCCAAGATGCGCTCCGCCCACACCAACGA  
CGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCACCGAGTCCATCG  
TGATCTGGGGCAAGACCCCCAAGTTCCGCCCTGCCCATCCAGAAGGAGACC  
TGGGAGACCTGGTGGACCGAGTACTGGCAGGCCACCTGGATTCCCGAGTG  
GGAGTTCGTGAACACCCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGA  
AGGAGCCCATCGCCGGCGCCGAGACCTTCTACGTGGACGGCGCCGCCAAC  
CGCGAGACCAAGCTGGGCAAGGCCGGCTACGTGACCGACCGCGGCCGCCA  
GAAGGTGGTGTCCCTGACCGAGACCACCAACCAGAAAACCGAGCTGCAGG  
CCATCCACCTGGCCCTGCAGGACTCCGGCTCCGAGGTGAACATCGTGACC  
GACTCCCAGTACGCCCTGGGCATCATCCAGGCCAGCCCGACAAGTCCGA  
GTCCGAGCTGTGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAAGG  
TGTACCTGTCTGGGTGCCCGCCACAAGGGCATCGGCGGCAACGAGCAG  
GTGGACAAGCTGGTGTCCACCGGCATCCGCAAGGTGCTGTTCTGGACGG  
CATCGACAAGGCCAGGAGGAGCACGAGAAGTACCACTCCAAGTGGCGCG  
CCATGGCCTCCGACTTCAACCTGCCCCCATCGTGCCCAAGGAGATCGTG  
GCCTCCTGCGACAAGTGCCAGCTGAAGGGCGAGGCCATGCACGGCCAGGT  
GGACTGCTCCCCCGGCATCTGGCAGCTGGACTGCACCCACCTGGAGGGCA  
AGATCATCCTGGTGGCCGTGCACGTGGCCTCCGGCTACATCGAGGCCGAG  
GTGATCCCCGCGGAGACCGGCCAGGAGACCGCCTACTTCATCCTGAAGCT  
GGCCGGCCGCTGGCCCGTGAAGGTGATCCACACCGACAACGGCTCCAAC  
TCACCTCCGCGCCGTGAAGGCCGCTGCTGGTGGGCCGGCATCCAGCAG  
GAGTTCGGCATCCCCACACCCCCAGTCCCAGGGCGTGGTGGAGTCCAT  
GAACAAGGAGCTGAAGAAGATCATCGGCCAGGTGCGCGACCAGGCCGAGC  
ACCTCAAGACCGCCGTGCAGATGGCCGTGTTTCATCCACAACCTCAAGCGC  
AAGGGCGGCATCGGCGGCTACTCCGCCGGCGAGCGCATCATCGACATCAT  
CGCCACCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCACCAAGATCC  
AGAACTTCCGCGTGTACTACCGCGACTCCCGCGACCCCCATCTGGAAGGGC  
CCCCCAAGCTGTGTGGAAGGGCGAGGGCGCCGTGGTGGTATCCAGGACAA  
CTCCGACATCAAGGTGGTGCCCCCGCGCAAGGCCAAGATCATCCGCGACT  
ACGGCAAGCAGATGGCCGGCGACGACTGCGTGGCCGGCCCGCCAGGACGAG  
GACTAA



29/178

## Fig. 19C

M.con.nef (group M consensus nef. Identical amino acid sequence to that in the public domain)

GCCGCCGCCATGGGCGGCAAGTGGTCCAAGTCCTCATCGTGGGCTGGCC  
CGCCGTGCGCGAGCGCATCCGCCGCACCCACCCCGCCGAGGGCGTGG  
GCGCCGTGTC CAGGACCTGGACAAGCACGCGCCATCACCTCTCCAA  
ACCGCCGCCAA CAACCCCGACTGCGCTGGCTGGAGGCCAGGAGAGGA  
GGAGGAGGTGGGCTTC CCGTGCGCC CAGGTGCCCTGCGCCCATGA  
CCTACAAGGCCGCTGGACCTGTC CCACTTCCTGAAGGAGAAGGCGGC  
CTGGAGGGCCTGATCTACTCCAAGAAGCGC CAGGAGATCCTGGACCTGTG  
GGTGTA CCA CACCCAGGGCTACTTC CCGACTGGCAGAAC TACACCCCG  
GCCCCGGCATCGCTA CCCCCTGACCTTCGGCTGGTGCTTCAAGCTGGTG  
CCCGTGGACCCCGAGGAGGTGGAGGAGGCCAACGAGGGCGAGAACAACTC  
CCTGCTGCA C CCAATGTGCCAGCACGGCATGGAGGACGAGGAGCGCGAGG  
TGCTGATGTGGAAGTTCGACTCCCGCTGGCCCTGCGCCACATCGCCCGC  
GAGCTGCACC CCGAGTACTACAAGGACTGCTAA

## Fig. 19D

C.con.pol.nuc

GCCGCCGCCATGCCCCAGATCACCTGTGGCAGCGCCCCCTGGTGTCCAT  
CAAGGTGGGCGGCCAGATCAAGGAGGCCCTGCTGGCCACCGGCGCCGACG  
ACACCGTGCTGGAGGAGATCAACCTGCCCGGCAAGTGGAAGCCCAAGATG  
ATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCT  
GATCGAGATCTGCGGCAAGAAGGCATCGGCACCGTGCTGGTGGGCCCCA  
CCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGCACC  
CTGAACCTCCCCATCTCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCC  
CGGCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCCTGACCGAGGAGAAGA  
TCAAGGCCCTGACCGCCATCTGCGAGGAGATGGAGAAGGAGGGCAAGATC  
ACCAAGATCGGCCCCGAGAACCCCTACAACACCCCCGTGTTCCGCATCAA  
GAAGAAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA  
ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCC  
GCCGGCCTGAAGAAGAAGAAGTCCGTGACCGTGCTGGACGTGGGCGACGC  
CTACTTCTCCGTGCCCCCTGGACGAGGGCTTCCGCAAGTACACCGCCTTCA  
CCATCCCCCTCCATCAACAACGAGACCCCCGGCATCCGCTACCAGTACAAC  
GTGCTGCCCCAGGGCTGGAAGGGCTCCCCCGCCATCTTCCAGTCTCCAT  
GACCAAGATCCTGGAGCCCTTCCGCGCCCGAGAACCCCGAGATCGTGATCT  
ACCAGTACATGGACGACCTGTACGTGGGCTCCGACCTGGAGATCGGCCAG  
CACCGCGCCAAGATCGAGGAGCTGCGCGAGCACCTGCTGAAGTGGGGCTT  
CACCACCCCCGACAAGAAGCACCAAGGAGCCCCCTTCTGTGGATGG  
GCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCCAGCTGCCC  
GAGAAGGACTCCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCT  
GAACTGGGCCTCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCA  
AGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCCTGACCGAG  
GAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGAAGGAGCCCGT  
GCACGGCGTGTA TACTACGACCCCTCCAAGGACCTGATCGCCGAGATCCAGA  
AGCAGGGCCACGACCAGTGGACCTACCAGATCTACCAGGAGCCCTTCAAG  
AACCTCAAGACCGGCAAGTACGCCAAGATGCGCACCGCCCAACCAACGA  
CGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGTCCATCG  
TGATCTGGGGCAAGACCCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACC  
TGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATTCCCGAGTG  
GGAGTTCTGTGAACACCCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGA  
AGGAGCCCATCGCCGGCGCGAGACCTTCTACGTGGACGGCGCCGCCAAC

CGGAGACCAAGATCGGCAAGCGGGCTACGTGACCGACCGCGGGCGCCA  
GAAGATCGTGTCCCTGACCCGAGACCAACACAGAAACCGAGCTGCAGG  
CCATCCAGCTGGCCCTGCAGACTCCGGCTCCGAGGTGAACATCGTGACC  
GACTCCAGTACGCCCTGGCATCATCCAGGCCAGCCCGACAAGTCCGA  
GTCCGAGCTGTTGAACACAGATCATCGAGCAGCTGATCAAGAAGGAGCGCG  
TGTAACCTGTCTGGGTGCCGCCCAAGGGCATCGCGGCAACGAGCAG  
GTGGACAAGCTGGTGTCTCCGGCATCCGCAAGTGTCTGTTCTTGGACGG  
CATCGACAAGGCCAGGAGGACGAGAAAGTACCACTCCAACCTGGCGCG  
CCATGGCTCCGAGTTCAACCTGCCCTCCATCGTGGCCAAGGATCGTG  
GCCTCTGCGACAAGTGCAGCTGAAGGGGAGGCCATGCACGGCCAGGT  
GGACTGCTCCCCGGCATCTGGCAGCTGGAATGCACCCACCTGGAGGGCA  
AGATCATCTGTGGCGTGCACGTGGCTCCGGCTACATCGAGGCCGAG  
GTGATCCCCCGGAGACCGGCCAGGAGACCGCTACTTCATCTGAAAGCT  
GGCGGCCGTGGCCCGTGAAGGTGATCCACACCGACAACGGCTCCAAC  
TCACCTCCGCCCGGTGAAGCGCTGCTGTGGCGCGCATCCAGCAG  
GAGTTCGGCATCCCCTAACACCCAGTCCCAGGGCGTGTGGAGTCCAT  
GAACAAGAGCTGAAGAAGATCATCGCCAGTGCAGCCAGGCCGAGC  
ACCTCAAGACCGCGGTGCAGATGGCCGTGTTCATCCACAACCTCAAGCGC  
AAGGGCGCATCGCGGTACTCCGCCGGGAGCGCATCATCGACATCAT  
CGCCACCGACATCCAGACCAAGGAGTGCAGAAAGCAGATCATCAAGATCC  
AGAACTTCCGCGTGTAACCGGACTCCCGGAGCCCACTGGAAGGGC  
CCCCCAAGCTGTGTGAAGGGGAGGGCGCGGTGTGATCCAGGACAA  
CTCCGACATCAAGGTGTGCCCCCGCAAGGCCAAGATCATCAAGGACT  
ACGGCAAGCAGATGGCCGGCGGCGACTGCGTGGCGCGCCCGCAGGACGAG  
GACTAA

Fig. 19D (continued)

M.con.gag (group M consensus gag)

MGARASVLSGKLDWEKIRLRPGGKKYRLKHLVWASRELERFALNPGLLETSEG CKQIIGQLQPA  
LQTGSEELRSLYNTVATLYCVHQRIEVKDTKEALEKIEEEQNKSQQTQAAADKGNSSKVSQNYPIVQN  
LQGQMVHQAI SPRTLNAWKVIEEKAFSPEVIPMFSALSEGATPDINTMLNTVGGHQAAQMMLKDTINE  
EAAEWDRLHPVHAGPI PPGOMREPRGSDIAGTSTLQEQIAWMTSNPPI PVGEIYKRWI ILGLNKIVRM  
SPVSILDIRQGPKEFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQNANPDCKTILKALPGATLEEM  
TACQGVGGPGHKARVLAEMSQVTNAAIMQRGNFKGORRI IKFNCGKEGHIARNCRAPRKKGCWKCGK  
EGHQMKDCTERQANFLGKIWPSNKGPRGNFLQSRPEPTAPPAESFGFGEIITPSPKQEPKDEPPLTSLK  
SLFGNDPLSQ

Fig. 19E